DSE-316/616 : Deep Learning - Assignment1

Name:	**Alli Khadga Jyoth**
Registration No./Roll No.:	19024
Institute/University Name:	IISER Bhopal
Program/Stream:	DSE
Date of Submission:	October 4, 2022

Question II

0.1 Introduction

There are 3 CNN architectures which are compared to each other using the CIFAR10 dataset. The architectures/models are as follows:

(A) Conv-Pool-Conv-Pool-Conv-Pool-FC

(B) Conv-Conv-Pool-Conv-Conv-Pool-FC

(C) Conv-Pool-Conv-Pool-Conv-Pool-FC-FC

From here on I will refer to them as model1, model2 and model3 respectively.

0.2 model-1

model1 contains three alternating conv and pool layers with a fully connected layer at the end. The accuracy and losses can be seen form the fig 1. Best accuracy: 66.9

0.3 Model-2

model2 contains 4 conv layers with two con layers stacked together. The accuracy and losses of model 2 can be found at fig 2. Best accuracy : 70.7

0.4 Model-3

model3 contains 3 conv-pool layers with two fully connected layers at the end. The accuracy and Loss plots can be given in fig 3. Best accuracy : 68.17
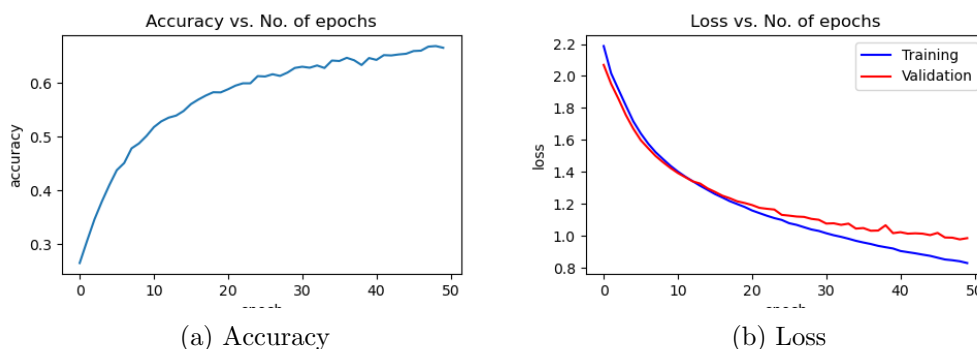

(a) Accuracy

(b) Loss

Figure 1: Model-1 accuracy and Loss

(a) Accuracy

(b) Loss

Figure 2: Model-2 accuracy and Loss

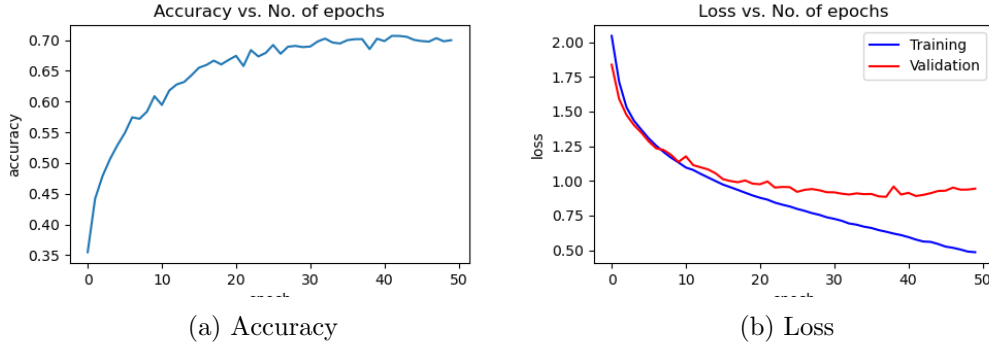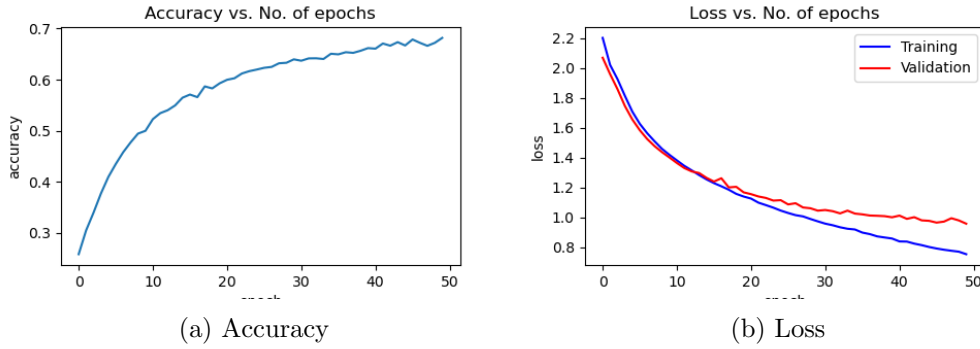(a) Accuracy

(b) Loss

Figure 3: Model-2 accuracy and Loss

0.4.1 Parameters

These are the initial parameters which the three models were trained using.

Loss function	Cross entropy
Optimizer	SGD with momentum
Learning rate	0.001
Epochs	50
batch size	256

a Hyperparameters

Pooling	Max pool
Pool size	3
Pool stride	2

b model Parameters

Table 1: Hyperparameters and Model architecture parameters

Table 2: Model Size in number of Params

Pooling Size	Stride	Model1	Model2	Model3
2	1	770,378	1,627,466	34,767,434
2	2	98,378	452,426	361,034
3	1	605,258	1,412,426	26,313,290
3	2	94,538	429,386	164,426
4	1	463,178	1,217,866	19,038,794

A:How does changing the network size change the accuracy?

From comparing the testing accuracy of three models, it can be seen that the model2 performs the best out of the three models given a preset parameters. This can be explained by the use of more conv layers in the model2, therefore the model2 was able to learn more of the dataset features more than

2

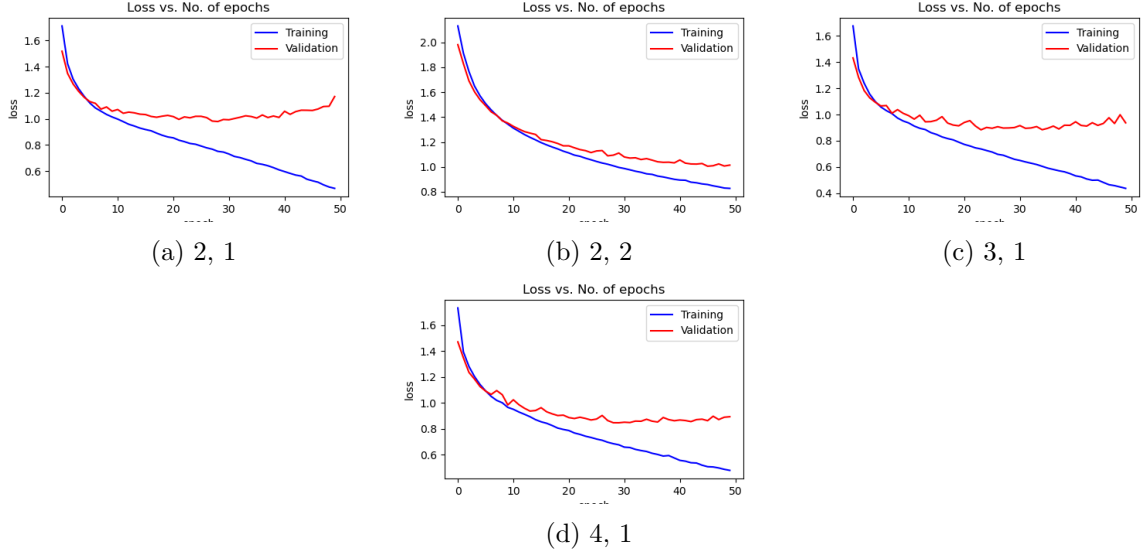(a) 2, 1   (b) 2, 2   (c) 3, 1

(d) 4, 1

Figure 4: Model-1 Loss (Pool Size, Stride)

the other two models. But comparing the training loss and validation loss, it is seen that the model2 suffers from over-fitting. This behaviour is can also be explained by the use of more conv layers. The model was able to learn the important features quickly, because of more conv layers, therefore the loss decreases quickly. Then the model tries to memorise the training data, decreasing the training loss, but increasing its validation/testing loss.

B:Experiment with different sizes of pooling and do a detailed analysis of pooling size on the network

From the Table 3, it is evident that pooling Size and Stride has a significant impact on the Model accuracy. From the data, for a given model the accuracy generally increases with increase in Pooling Size and for a given Pooling Size the accuracy decreases with increase in its stride. Therefore the maximum accuracy is achieved in every model with a Pooling Size of 4 and a stride of 1.

Table 3: Model Accuracy[1]w.r.to Pooling Size and Stride

Pooling Size	Stride	Model1	Model2	Model3
2	1	67.1	66.7	67.7
2	2	65.72	67.21	66.13
3	1	71.72	71.5	71.9
3	2	66.9	70.7	68.17
4	1	**72.43**	**72.6**	**72.41**

C. How the presence of one or more fully connected layers changes the accuracy

We can compare Model1 and Model3 directly, since they both have same architecture apart from an extra FC layer at the end in model3. From the table 3, comparing the accuracies of model1 and model3 at various Pooling Sizes and Strides, we can conclude that model1 and model3 perform almost similar. Although model3 is less susceptible to stride changes within a given Pooling Size. This is directly seen when, we fix a pooling size (say 3), and we change the Stride. Both models perform the same at stride 1, but when we change the stride to 2, model1 has a much more drop in performance compared to model3.
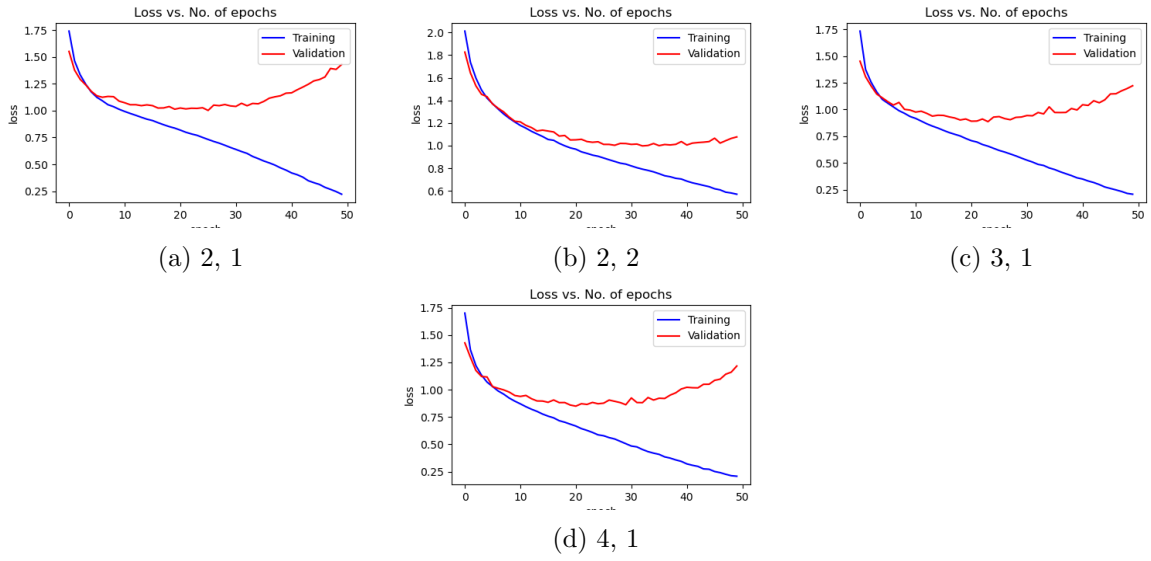
---

[1]Model Best Accuracy

(a) 2, 1 (b) 2, 2 (c) 3, 1

(d) 4, 1

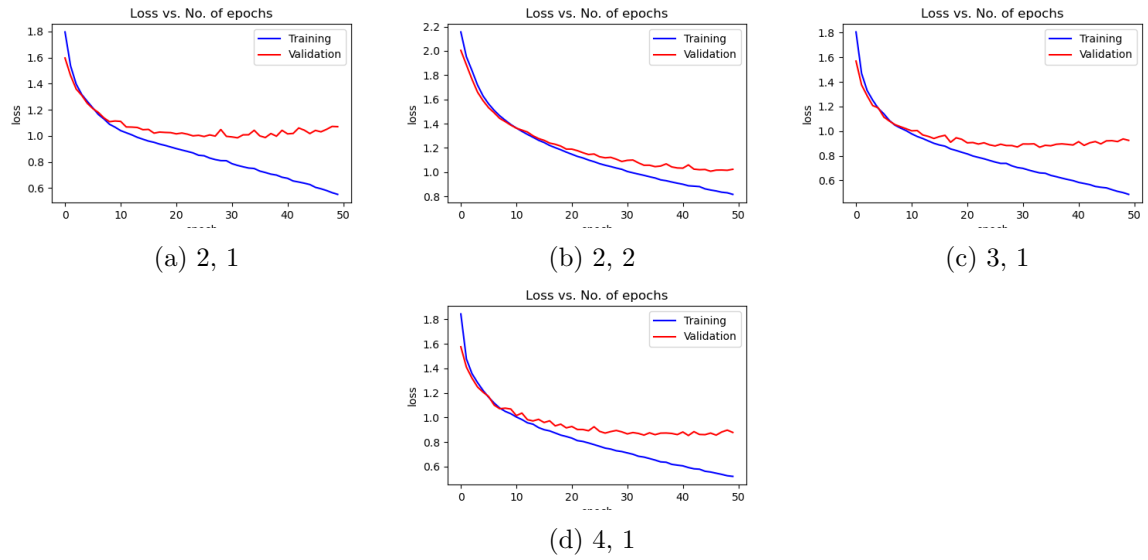Figure 5: Model-2 Loss (Pool Size, Stride)

(a) 2, 1 (b) 2, 2 (c) 3, 1

(d) 4, 1

Figure 6: Model-3 Loss (Pool Size, Stride)